In situ localization of the transcripts of a homeobox gene in the honeybee *Apis mellifera* L. (Hymenoptera)

Richard Fleig¹, Uwe Walldorf¹, Walter Jakob Gehring¹, and Klaus Sander²

¹ Biozentrum, Klingenbergstrasse 70, CH-4056 Basel, Switzerland

² Institut für Biologie I (Zoologie), Albertstrasse 21a, D-7800 Freiburg, Federal Republic of Germany

Summary. We have isolated and characterized a homeoboxcontaining gene from the honeybee Apis mellifera. Its homeobox region shows a high degree of sequence similarity to the homeobox of the Drosophila gene Deformed (Dfd). At the DNA level 82% of the basepairs are the same, whereas the putative amino acid sequences are identical between the bee and the fruitfly genes. Similarity is also present 5' and 3' to the homeobox. Using this isolate as a probe we have performed in situ hybridization on sections from blastoderm-stage embryos of the honeybee Apis mellifera. In early blastoderm stages we found a rather irregular pattern of labelled nuclei. In middle stages we found silver grains over each nucleus and also over the cytoplasm in a belt of blastoderm cells in the prospective gnathal region. These results indicate that the Deformed genes from honeybee and fruitfly are homologous both with respect to their DNA sequence and their spatial and temporal pattern of expression during embryogenesis.

Key words: Apis mellifera – Homeobox genes – Dfd – In situ hybridization – Blastoderm

Introduction

A common feature of many metazoa is the subdivision of their body into homologous segments. In Drosophila melanogaster it has been shown that two classes of genes are required for establishing the characteristic segmentation pattern. Segmentation genes establish the basic metameric units and homeotic genes determine the identity of each unit (Nüsslein-Volhard et al. 1982). Many of these genes share a similar 180-bp nucleotide sequence, the homeobox (McGinnis et al. 1984a; Scott and Weiner 1984). Sequence similarity of homeobox proteins with yeast mating-type sequences and with several bacterial DNA-binding proteins (Sheperd et al. 1984; Laughon and Scott 1984) has led to the proposal that the homeo domain may bind to DNA. DNA binding studies (Desplan et al. 1985) have further supported this hypothesis. The general significance of the homeobox was shown by the discovery of these sequences in vertebrates; thus far homeobox sequences have been isolated from frog, rat, mouse and man (Carrasco et al. 1984;

Offprint requests to: K. Sander

Müller et al. 1984; Falzon et al. 1987; McGinnis et al. 1984c; Colberg-Poley et al. 1985; Levine et al. 1984).

Using Southern blot analysis and molecular cloning we have recently shown that homeobox sequences are also present in the honeybee Apis mellifera (unpublished results). Due to its large size (1.7 mm) and slower development, the honeybee egg offers some advantages to monitor early patterns of gene expression more precisely. Development until hatching takes about 70 h at 35° C (Nelson 1915; Schnetter 1935; Du Praw 1967; Fleig and Sander 1986). The blastoderm stage begins about 8 h after egg deposition and lasts until 33 h, when gastrulation starts; between 10 h and 34 h no mitoses can be observed in the scanning electron microscope (SEM) (Fleig and Sander 1985). Thus the absolute, as well as the relative, duration of the blastoderm period is long when compared to that of Drosophila melanogaster (Turner and Mahowald 1976; Campos-Ortega and Hartenstein 1985) and many other insects.

One of the homeoboxes we have found shows strong sequence similarity to the homeobox of the *Deformed (Dfd)* gene of *Drosophila melanogaster*. In the fruitfly, *Dfd* is involved in establishing head structures (Morgan et al. 1925; Regulski et al. 1987). In situ hybridizations at blastoderm and germ-band stages of *Drosophila* reveal a belt of label in an area which approximately contains the anlagen of the maxillary and the mandibular segments (McGinnis et al. 1984a; Harding et al. 1985; Gehring 1987; Chadwick and McGinnis 1987; Martinez-Arias et al. 1987). Our in situ hybridizations at mid-blastoderm stages of the honeybee show a similar belt of cytoplasmic transcript accumulation. This indicates homology of the *Drosophila* and the *Apis* genes, not only at the DNA level, but also with respect to time and region of gene activity during embryogenesis.

Material and methods

Materials. Honeybees (*Apis mellifera*) and honeybee eggs were collected from bee colonies kept at the University of Freiburg, Germany. Radionucleotides were purchased from Amersham and all enzymes from Biofinex or Boehringer.

General methods. Preparation of genomic DNA was as described earlier (Walldorf et al. 1984). Restriction endonuclease digestions, gel electrophoresis of DNA, labelling of DNA fragments, Southern transfer experiments and the construction of a genomic library were performed as described by Maniatis et al. (1982).

Library screening. A library of *Sau*3A partially digested genomic *Apis mellifera* DNA cloned into the *Bam*HI site of the EMBL4 phage lambda vector (Frischauf et al. 1983) has been used for screening. 1.5×10^5 phages (5 genome equivalents) were screened under low stringency conditions at 37° C. The hybridization buffer consisted of 43% deionized formamide, $5 \times SSC$ ($1 \times SSC$ is 0.15 *M* NaCl, 15 m*M* sodium citrate), $4 \times$ Denhardt's solution, 0.1% SDS, 0.1% sodiumpyrophosphate, 20 µg/ml tRNA and 50 µg/ml heparin. The filters were washed twice for 30 min each at 50° C in $2 \times SSC$.

DNA sequencing. All DNA sequencing procedures were carried out using the M13 cloning (Messing and Vieira 1982) and chain termination sequencing (Sanger et al. 1977) methods. Both strands of the DNA were sequenced.

In situ hybridization. Eggs were collected every 2 h from a honeybee queen (*Apis mellifera carnica*) encaged on a single empty honeycomb. They were incubated in darkness at 35° C and high humidity up to the appropriate age and developmental stage (Fleig and Sander 1986). After dechorionation with sodium hypochlorite the eggs were fixed for 1 h with glutaraldehyde in the heptane phase of a glutaraldehyde/*n*-heptane mixture (Zalokar and Erk 1977). Cryosectioning and in situ hybridization of the sections was performed as described for *Drosophila* by Hafen et al. (1983). As hybridization probe we used the 0.34-kb *SacI-SalI* fragment (Fig. 2). It includes the greater part of the homeobox and the M repeat 3' of the box (Fig. 3). Labelling of the probe with tritium was done by nick translation. Time of exposure was 5 weeks in a dry chamber at 4° C.

Results

Using Southern blot analysis we have previously shown that at least four EcoRI fragments of honeybee DNA hybridize with the Drosophila homeobox probes Antp, ftz, Scr and Dfd (Walldorf et al. unpublished results). An example of such a genomic Southern blot is shown in Fig. 1. Here the homeobox sequence of the Drosophila gene Dfd was used as a probe. In the lane with genomic Drosophila DNA several bands are visible, the strongest being the Deformed fragment itself, whereas the others represent cross-hybridizing fragments of other homeobox containing genes. Using genomic honeybee DNA the strongest band corresponds to a size of 4.2 kb, a 1.5-kb band being also quite strong and bands of 5.5 kb and 9.0 kb are less intense. Since the strongest band always has the highest sequence similarity to the probe, we considered the 4.2-kb fragment to be closely related to the Drosophila Dfd homeobox, whereas the other bands represent other homeobox-containing genes which are more diverged (Walldorf et al. unpublished results). For further analysis we used the lambda phage B4, which was isolated during a screen of a genomic honeybee library under low stringency conditions with the Antp probe (Walldorf et al. unpublished results). This clone contains the 4.2-kb *Eco*RI fragment mentioned above, which hybridizes with both Antp and with Dfd. The fragment was subcloned into pUC8 and a restriction map established (Fig. 2). We found the homeobox sequence similarity to be located within a 500-bp Cla-SalI fragment whose DNA sequence was determined.



Fig. 1. Homeobox similarity in the honeybee genomic DNA. Genomic DNAs were digested with *Eco*RI and hybridized under low stringency conditions with a *Drosophila Dfd* homeobox probe (*Dfd* 250-bp *Hpa*II fragment). Lane 1, 2.5 μ g *Drosophila* DNA; lane 2, 7.5 μ g *Apis mellifera* DNA. Size markers are lambda *Hind*III fragments

pH42



Fig. 2. Restriction map of clone pH42. The *black box* identifies the region containing homeobox similarity, 5' and 3' is indicated. Restriction enzymes: B, *Bam*HI; Bg, *BgI*II; C, *Cla*; E, *Eco*RI; H, *Hind*III; S, *SaI*I; Sa, *SacI. Arrows* below the map represent sequence determinations

The fragment has a complete homeobox sequence which is 82% identical to a cDNA sequence of the *Drosophila* Dfd gene's homeobox (Fig. 3). More striking is the fact that the homeoboxes of the Dfd gene and of the honeybee gene (termed H42) code for an identical amino acid sequence; in addition, sequence similarity extends nine amino acids beyond both the 5' and 3' ends of the box (Fig. 3). This translates into a total of 78 identical amino acids in Dfd and H42. 5' of the homeobox the sequence similarity

	CGT	GCC	CGA	тст	CTG	TTT	CGA	TTT	ATT	тса	CG GCG	AAC AAC	GGG GGC	TCT TCG	TAC TAC	CAG CAG	CCG CCC	GGA GGG	ATG ATG	GAG GAG
H42												ASN	GLY	SER	TYR	GLN	PRO	GLY	MET	GLU
Dfd	1 PRO	LYS	ARG	GLN	ARG	THR	ALA	TYR	THR	ARG	ніѕ	GLN	ILE	LEU	GLU	LEU	GLU	LYS	GLU	PHE
	CCA CCG	ааа Ааа	CGC CGG	CAA CAG	CGC AGG	ACC ACC	GCC GCG	ТАС ТАС	ACA ACG	CGC AGG	CAT CAC	CAG CAG	ATC ATC	CTG CTC	GAA GAG	CTG CTC	GAA GAG	AAG AAG	GAG GAG	TTC TTC
H42	PRO	LYS	ARG	GLN	ARG	THR	ALA	TYR	THR	ARG	HIS	GLN	ILE	LEU	GLU	LEU	GLU	LYS	GLU	PHE
Dfd	21 HIS	TYR	ASN	ARG	TYR	LEU	THR	ARG	ARG	ARG	ARG	ILE	GLU	ILE	ALA	HIS	THR	LEU	VAL	LEU
	CAC CAT	TAC TAC	ААС ААТ	CGC AGG	TAC TAC	CTG CTG	ACG ACG	CGT AGA	CGG CGG	CGG CGG	CGC CGT	АТС АТС	GAG GAG	АТТ АТС	GCC GCT	CAT CAC	ACG ACC	TTA CTG	GTT GTC	CTC CTC
H42	HIS	TYR	ASN	ARG	TYR	LEU	THR	ARG	ARG	ARG	ARG	ILE	GLU	ILE	ALA	HIS	THR	LEU	VAL	LEU
Dfd	41 SER	GLU	ARG	GLN	ILE	LYS	ILE	TRP	PHE	GLN	ASN	ARG	ARG	MET	LYS	TRP	LYS	LYS	ASP	ASN
	TCG TCC	GAG GAG	CGG CGG	CAG CAG	ATC ATC	AAG AAG	АТС АТС	TGG TGG	TTC TTC	CAG CAG	AAC AAC	AGG CGT	CGC CGG	ATG ATG	AAG AAG	TGG TGG	AAG AAG	AAG AAG	GAC GAC	AAC AAC
Н42	SER	GLU	ARG	GLN	ILE	LYS	ILE	TRP	PHE	GLN	ASN	ARG	ARG	MET	LYS	TRP	l,ys	LYS	ASP	ASN
Dfd	61 Lys	LEU	PRO	ASN	THR	LYS	ASN	VAL	ARG	LYS	LYS	THR	VAL	ASP	ALA	ASN	GLY	ASN	PRO	THR
Dfđ	61 LYS AAG AAG	LEU CTG CTG	PRO CCC CCC	ASN AAC AAC	THR ACC ACG	LYS AAG AAG	ASN AAC AAC	VAL GTG GTG	ARG CGC AGG	LYS AAG CGG	LYS AAG AAG	THR ACG AAC	VAL GTG GGG	ASP GAC GGC	ALA GCC CAG	ASN AAC GCC	GLY GGC GCA	ASN AAC CGT	PRO CCA CCG	THR ACA CCG
Dfđ H42	61 LYS AAG AAG LYS	LEU CTG CTG LEU	PRO CCC CCC PRO	ASN AAC AAC ASN	THR ACC ACG THR	LYS AAG AAG LYS	ASN AAC AAC ASN	VAL GTG GTG VAL	ARG CGC AGG ARG	LYS AAG CGG ARG	LYS AAG AAG LYS	THR ACG AAC ASN	VAL GTG GGG GLY	ASP GAC GGC GLY	ALA GCC CAG GLN	ASN AAC GCC ALA	GLY GGC GCA ALA	ASN AAC CGT ARG	PRO CCA CCG PRO	THR ACA CCG PRO
Dfđ H42 Dfđ	61 LYS AAG AAG LYS 81 PRO	LEU CTG CTG LEU VAL	PRO CCC CCC PRO	ASN AAC AAC ASN	THR ACC ACG THR	LYS AAG AAG LYS PRO	ASN AAC AAC ASN	VAL GTG GTG VAL	ARG CGC AGG ARG ARG	LYS AAG CGG ARG ALA	LYS AAG AAG LYS ALA	thr acg aac asn ser	VAL GTG GGG GLY LYS	ASP GAC GGC GLY LYS	ALA GCC CAG GLN	ASN AAC GCC ALA GLN	GLY GGC GCA ALA GLN	ASN AAC CGT ARG ALA	PRO CCA CCG PRO GLN	THR ACA CCG PRO GLN
Dfd H42 Dfd	61 LYS AAG AAG LYS 81 PRO CCG GCA	LEU CTG CTG LEU VAL GTA AGT	PRO CCC PRO ALA GCG CGT	ASN AAC AAC ASN LYS AAG CCG	THR ACC ACG THR LYS AAA GCA	LYS AAG AAG LYS PRO CCC AGG	ASN AAC AAC ASN THR ACC GGG	VAL GTG GTG VAL LYS AAG GGG	ARG CGC AGG ARG ARG CGG	LYS AAG CGG ARG ALA GCC CGT	LYS AAG AAG LYS ALA GCC CGA	THR ACG AAC ASN SER TCC GCA	VAL GTG GGG GLY LYS AAA GGG	ASP GAC GGC GLY LYS AAG GGG	ALA GCC CAG GLN GLN CAG GGC	ASN AAC GCC ALA GLN CAG CGA	GLY GGC GCA ALA GLN CAA AAC	ASN AAC CGT ARG ALA GCG TGC	PRO CCA CCG PRO GLN CAG GGC	THR ACA CCG PRO GLN CAG CGG
Dfd H42 Dfd H42	61 LYS AAG AAG LYS 81 PRO CCG GCA ALA	LEU CTG LEU VAL GTA AGT SER	PRO CCC PRO ALA GCG CGT ARG	ASN AAC AAC ASN LYS AAG CCG PRO	THR ACC THR LYS AAA GCA ALA	LYS AAG AAG LYS PRO CCC AGG ARG	ASN AAC AAC ASN THR ACC GGG GLY	VAL GTG GTG VAL LYS AAG GGG GLY	ARG CGC AGG ARG CGG CGT ARG	LYS AAG CGG ARG ALA GCC CGT ARG	LYS AAG AAG LYS ALA GCC CGA ARG	THR ACG AAC ASN SER TCC GCA ALA	VAL GTG GGG GLY LYS AAA GGG GLY	ASP GAC GGC GLY LYS AAG GGG GLY	ALA GCC GLN GLN CAG GGC GLY	ASN AAC GCC ALA GLN CAG CGA ARG	GLY GGC GCA ALA GLN CAA AAC ASN	ASN AAC CGT ARG ALA GCG TGC CYS	PRO CCA CCG PRO GLN CAG GGC GLY	THR ACA CCG PRO GLN CAG CGG ARG
Dfđ H42 Dfđ H42 Dfđ	61 LYS AAG AAG LYS 81 PRO CCG GCA ALA 101 GLN	LEU CTG CTG LEU VAL GTA AGT SER GLN	PRO CCC PRO ALA GCG CGT ARG	ASN AAC AAC LYS AAG CCG PRO SER	THR ACC ACG THR LYS AAA GCA ALA	LYS AAG AAG LYS PRO CCC AGG ARG GLN	ASN AAC AAC ASN THR ACC GGG GLY GLN	VAL GTG GTG VAL LYS AAG GGG GLY	ARG CGC AGG ARG CGG CGT ARG	LYS AAG CGG ARG ALA GCC CGT ARG	LYS AAG AAG LYS ALA GCC CGA ARG	THR ACG AAC ASN SER TCC GCA ALA THR	VAL GTG GGG GLY LYS AAA GGG GLY PRO	ASP GAC GGC GLY LYS AAG GGG GLY VAL	ALA GCC CAG GLN CAG GCC CAG GCZ GLY MET	ASN AAC GCC ALA GLN CAG CGA ARG ASN	GLY GGC GCA ALA GLN CAA AAC ASN GLU	ASN AAC CGT ARG ALA GCG TGC CYS	PRO CCA CCG PRO GLN CAG GCC GLY	THR ACA CCG PRO GLN CAG CGG ARG ARG
Dfd H42 Dfd H42 Dfd	61 LYS AAG AAG LYS 81 PRO CCG GCA ALA 101 GLN CAG GAA	LEU CTG CTG LEU VAL GTA AGT SER GLN CAG CGG	PRO CCC PRO ALA GCG CGT ARG GLN CAG GAA	ASN AAC AAC ASN LYS AAG CCG PRO SER TCG CAA	THR ACC ACG THR LYS AAA GCA ALA CAG CAA	LYS AAG AAG LYS PRO CCC AGG ARG CAG CAG	ASN AAC AAC ASN THR ACC GGG GLY GLN CAG CCA	VAL GTG GTG VAL LYS AAG GGG GLY CAG GAC	ARG CGC AGG ARG CGG CGT ARG THR ACG CAG	LYS AAG CGG ARG ALA GCC CGT ARG GLN CAG GAA	LYS AAG AAG LYS ALA GCC CGA ARG GLN CAG GAA	THR ACG AAC ASN SER TCC GCA ALA THR ACT CAA	VAL GTG GGG GLY LYS AAA GGG GLY PRO CCG CAA	ASP GGC GLY LYS AAG GGG GLY VAL GTG CAA	ALA GCC CAG GLN CAG GCC GLY MET ATG CAA	ASN AAC GCC ALA CAG CAA ARG ASN AAT CAA	GLY GGC GCA ALA GLN CAA AAC ASN GLU GAG CTC	ASN AAC CGT ARG ALA GCG TGC CYS CYS TGC GCC	PRO CCA PRO GLN CAG GGC GLY ILE ATT GTC	THR ACA CCG PRO CAG CGG ARG ARG CGT GAT
Dfd H42 Dfd H42 Dfd H42	61 LYS AAG AAG B1 PRO CCG GCA ALA 101 GLN CAG GAA GLU	LEU CTG CTG LEU VAL GTA AGT SER GLN CAG CGG ARG	PRO CCC PRO ALA GCG CGT ARG GLN CAG GAA GLU	ASN AAC AAC ASN LYS AAG CCG PRO SER TCG CAA GLN	THR ACC ACG THR LYS AAA GCA ALA CAG CAG CAA GLN	LYS AAG AAG LYS PRO CCC AGG ARG CAG CAG CAG GLN	ASN AAC AAC ASN THR ACC GGG GLY CAG CCA PRO	VAL GTG GTG LYS AAG GGG GLY CAG GAC ASP	ARG CGC AGG ARG CGT ARG CGT ARG CGT ARG CGT ARG CGT	LYS AAG CGG ARG ALA GCC CGT ARG GLN CAG GAA GLU	LYS AAG AAG LYS ALA GCC CGA ARG CAG GAA GLU	THR ACG AAC ASN SER TCC GCA ALA THR ACT CAA GLN	VAL GTG GGG LYS AAA GGG GLY PRO CCG CAA GLN	ASP GAC GC GLY LYS AAG GGG GLY VAL GTG CAA GLN	ALA GCC CAG GLN CAG GGC GLY MET ATG CAA GLN	ASN AAC GCC ALA GLN CAG CGA ARG ASN AAT CAA GLN	GLY GGC GCA ALA CAA AAC ASN GLU GAG CTC LEU	ASN CGT ARG ALA GCG TGC CYS CYS TGC GCC ALA	PRO CCA CCG PRO CAG GGC GLY ILE ATT GTC VAL	THR ACA CCG PRO CAG CGG ARG ARG CGT GAT ASP
Dfd H42 Dfd H42 Dfd H42 Dfd	61 LYS AAG AAG E <u>LYS</u> 81 PRO CCG GCA ALA 101 <u>GLN</u> CAG GAA GLU 121 SER	LEU CTG CTG LEU VAL GTA AGT SER CAG CGG ARG	PRO CCC PRO ALA GCG CGT ARG GLN GLN GLU SER	ASN AAC AAC LYS AAG CCG PRO SER TCG CAA GLN	THR ACC ACG THR LYS AAA GCA ALA GLN GLU	LYS AAG LYS PRO CCC AGG ARG GLN CAG CAG GLN SER	ASN AAC ASN THR ACC GGG GLY CAG CCA PRO	VAL GTG GTG VAL LYS AAG GGG GLY CAG GAC ASP	ARG CGC AGG CGG CGT ARG CGG CAG CAG GLN ASP	LYS AAG CGG ARG ALA GCC CGT ARG GLN CAG GAA GLU VAL	LYS AAG LYS ALA GCC CGA ARG GLN GLU	THR ACG AAC ASN SER TCC GCA ALA THR CAA GLN	VAL GTG GGG GLY LYS AAA GGG GLY PRO CCG CAA GLN	ASP GAC GGC GLY LYS AAG GGG GLY VAL GTG CAA GLN	ALA GCC CAG GLN CAG GGC GLY MET ATG CAA GLN	ASN AAC GCC ALA CAG CGA ARG ASN AAT CAA GLN	GLY GGC GCA ALA CAA AAC CAA ASN GLU GAG CTC LEU	ASN AAC CGT ARG ALA GCG TGC CYS CYS TGC GCC ALA	PRO CCA CCG PRO CAG GGC GLY ILE ATT GTC VAL	THR ACA CCG PRO CAG CAG CAG CAG CAG CAG CAG CAG CAG ARG ASP
Dfd H42 Dfd H42 Dfd H42 Dfd	61 LYS AAG AAG AAG EYS 81 PRO CCG GCA ALA 101 GLN CAG GAA GLU 121 SER TCC GGC	LEU CTG CTG LEU VAL GTA AGT SER CAG CGG ARG ASP GAC CGA	PRO CCC PRO ALA GCG CGT ARG GLN GLN GLU SER AGT CCC	ASN AAC AAC LYS AAG CCG PRO SER TCG CAA GLN LEU	THR ACC ACG THR LYS AAA GCA ALA CAG CAG CAA GLN GLU GAG	LYS AAG AAG PRO CCC AGG ARG CAG CAG CAG CAG GLN SER AGT GAT	ASN AAC ASN THR ACC GGG GLY CAG CCA PRO ILE ATC GGA	VAL GTG GTG LYS AAG GGG GLY CAG GAC GLY GLY	ARG CGC AGG ARG CGG CGT ARG CGG CGT ARG CAG GLN ASP GAC GTC	LYS AAG CGG ARG CGT CGT ARG GLN CAG GLN CAG GLU VAL CGTC GAC	LYS AAG AAG LYS ALA GCC CGA ARG GLN CAG GAA GLU	THR ACG AAC ASN SER TCC GCA ALA THR ACT CAA GLN	VAL GTG GGG GLY LYS AAA GGG GLY PRO CCG CAA GLN	ASP GAC GGC GLY LYS AAG GGG GLY VAL GTG CAA GLN	ALA GCC CAG GLN CAG GGC GLY MET ATG CAA GLN	ASN AAC GCC ALA CAG CGA ARG ASN AAT CAA GLN	GLY GGC GCA ALA CAA AAC ASN GLU GAG CTC LEU	ASN AAC CGT ARG GCG TGC CYS CYS TGC GCC ALA	PRO CCA CCG PRO CAG GGC GLY ILE ATT GTC VAL	THR ACA CCG PRO CAG CGG ARG ARG CGT GAT ASP

Dfd

ASN GLY SER TYR GLN PRO GLY MET GLU

Fig. 3. DNA and protein sequence comparison of *Drosophila Dfd* and *Apis H42*. DNA and protein sequences of *H42* are aligned with the comparable region from *Dfd* (clone cDf 41, Regulski et al. 1987). Putative splice sites are indicated by *arrow*. The homeobox region is *boxed* and contiguous stretches of identical amino acids are underlined (*thick bar*). M repeat regions are also underlined (*thin bars*)

stops at the position where Dfd has an intron. Since we find a perfect consensus splice site at this position in H42, it is likely that it also has an intron 5' to the homeobox. Nine amino acids 3' to the homeobox the degree of sequence similarity drops although we still have an open reading frame in both clones. The only similarity in this region is due to M repeat sequences (McGinnis et al. 1984a; Wharton et al. 1985), located in both cases at a similar distance from the homeobox. Since the sequence comparisons favoured the hypothesis that the H42 honeybee gene is a true Dfd homologue, we tested this by comparing the temporal and spatial distribution of H42 transcripts in Apis mellifera with that of Deformed transcripts in Drosophila melanogaster.

In early blastoderm stage sections (17 h) clusters of silver grains are found only over the blastoderm nuclei (Fig. 4). In this stage the blastoderm temporarily seems to be bilayered, although the cells are still open towards the yolk sac (Nelson 1915; Fleig and Sander 1985). The volume of the bottle-shaped blastoderm cells is largely taken up by their nuclei. The label is not equal over the whole nucleus but is concentrated over a restricted region (Fig. 4). We find label also over the nuclei of vitellophages inside the yolk system. In the 18-h embryo the blastoderm nuclei form a monolayer again (Nelson 1915; Fleig and Sander 1985) (Fig. 5). In this stage, too, label is found only over restricted parts of the nuclei (Fig. 5), which are now located in the distal part towards the periphery of the blastoderm cells. Not all nuclei are labelled but we find a rather patchy pattern of labelled and unlabelled nuclei. During the next 4 h (no mitoses are observed in this stage) increasing numbers of nuclei show silver grain labelling.

In sections of mid-blastoderm stage (24–28 h) we find label over every nucleus in blastoderm cells and vitellophages (Fig. 6). As in younger stages, the labelled area is smaller than the nucleus itself. The nuclei of the extra-em-



Fig. 4a, b. Details of a section from an early blastoderm stage. a Anterior, b near posterior end. Label can be seen over the blastoderm nuclei, which are arranged in two layers. Label is restricted to a defined part of the nuclei (*arrows*). Age 17 h after egg laying, bar 10 μ m

Fig. 5. Section of an early blastoderm stage embryo. The nuclei have returned to a single-layered arrangement close to the periphery. Only nuclei are labelled, the blastodermal cytoplasm shows no label. Age 18 h after egg laying, bar 10 μ m



Fig. 7. Dark field photograph of a section from an intermediate blastoderm stage. The blastodermal cytoplasm is labelled between 60% and 75% egg length (posterior pole, to the *right*, is 0%). Label in posterior pole region is not cytoplasmatic. Age 26 h after egg laying, total length 1.4 mm

m

8

Fig. 8. Scanning electron micrograph of a gastrulation-stage embryo, lateral view, anterior to the *left*. Segmental grooves can be seen with differing clarity in head, thorax and abdomen. *m*, Maxillary segment; *ec*, ectoderm; *ms*, mesoderm; *ps*, preserosa. Age 35 h after egg laying, total length 1.2 mm

bryonic dorsal parts of the blastoderm (anlagen of dorsal strip, serosa and amnion) do not differ in labelling from those of the embryonic lateral and ventral parts which will give rise to ectoderm and mesoderm/endoderm, respectively. But now we find label also over the entire cytoplasm of all blastoderm cells in a belt between 60% and 75% egg length (posterior pole is 0%) (Fig. 7). Comparison with a gastrulating embryo (Fig. 8), which is already developing segmental grooves, shows that this region should contain the anlagen of the mandibular, maxillary and perhaps labial segments, or at least parts of them. The signal intensity over the cytoplasm and nuclei of the cells in the belt is much higher than over the nuclei outside the belt. Labelling is very strong in the middle of this belt and fades out towards its borders. The more strongly labelled central part of the belt (about 65%-70% egg length) is 15-20 cells wide, which equals about two segments (segment width is about 8-10 cells once grooves become visible). It may contain the complete maxillary and at least part of the mandibular segment anlage (compare Figs. 7, 8). The labelled belt runs all around the circumference of the blastoderm. It can be seen in both the embryonic and extra-embryonic regions and even in the mid-dorsal parts, where during the next hours the blastoderm will be interrupted temporarily (dorsal strip, Nelson 1915; Fleig and Sander 1985) (Fig. 7). La-

Fig. 6. Anterior part of section in Fig. 7. Every nucleus in both blastoderm and yolk is labelled (*arrows*). Anterior to the belt of cytoplasmic labelling (between *arrowheads*) an area of weaker labelling in the nuclei can be seen (bars)

bel over the nuclei located just outside the labelled belt seems to be weaker than in nuclei further away from the belt (Fig. 6).

In late blastoderm stage sections (age 30 h, about 3 h before the onset of gastrulation) we no longer observed label over the vitellophage nuclei. In this stage we find a pattern of narrow peripheral belts of labelling, but due to insufficient material we cannot as yet establish any spatial relationship between the labelled belts and individual segments or segmental grooves.

Discussion

Homeobox sequences of the Antp class show a high degree of conservation in different species, whereas flanking sequences are not conserved. Among vertebrates, sequence similarity outside the homeobox has been shown in a few cases (Boncinelli et al. 1985). The same is true for homeobox genes in different insects (Walldorf et al. unpublished results). Sequence similarity outside the homeoboxes between insects and vertebrates has up to now been described only for the Drosophila genes engrailed (en) (Joyner and Martin 1987) and Deformed (Dfd) (Regulski et al. 1987). Our sequence data demonstrate that the honeybee clone H42 is a Dfd homologue. As expected, homology is higher between the two insect species than between insects and vertebrates. A region coding for 78 amino-acids including the homeobox is absolutely identical in Dfd and H42. This is the first case of a homeobox being identical in two different species (of two different orders: Dipterans and Hymenopterans). The identical locations of splice sites 5' to the box and the presence of an M repeat 3' to the homeobox are further hints for a high degree of conservation of the gene structure in both species. Since we only have genomic clones we cannot compare *Dfd* to the whole honeybee homologue, but still this high degree of conservation is striking.

The pair of genes from the two species express a similar spatial and temporal pattern of activity, at least during the blastoderm stage. The basic body organisation of both *Drosophila melanogaster* and *Apis mellifera* is similar. Both are long germ developers, which rapidly subdivide the embryonic (ventral) part of the blastoderm into the complete set of future segmental units (see Krause 1939; Sander 1976). Their individual segments are easily homologized although specific differences exist; for instance, head involution and reduction of the last two abdominal segments are specialities of *Drosophila melanogaster* and other cyclorrhapic Dipterans, whereas a specific trait of *Apis mellifera* consists of the absence of germ-band elongation and retraction.

The labelled belt of *H42* expression in the mid-blastoderm stages of the honeybee embryo is positioned as *Dfd* expression is in the fruitfly during cellular blastoderm (McGinnis et al. 1984a; Harding et al. 1985; Gehring 1987; Chadwick and McGinnis 1987; Martinez-Arias et al. 1987). The comparison of SEM preparations of early gastrula stages with the belt of labelling in the blastoderm stage section shows that in the honeybee the belt is equivalent to about two segment anlagen. This coincides with the *Dfd* results in *Drosophila melanogaster* (McGinnis et al. 1984a; Harding et al. 1985; Gehring 1987; Chadwick and McGinnis 1987; Martinez-Arias et al. 1987).

Nuclei in all regions of the blastoderm, as well as the vitellophages, carry label for at least 7 h before we find

label over the cytoplasm in the blastoderm region between 60% and 75% egg length. These findings may indicate that the transcript is stored somehow for many hours in the nuclei, or even in a restricted part of them, before being released only in specific areas of the blastoderm; or a rapid degradation outside the nuclei prevents higher accumulation of the mRNA in the cytoplasm. In *Drosophila* no such nuclear storage of *Dfd* transcripts is reported. It might have been overlooked because of the short duration of the blastoderm stage, but in the case of the *Drosophila* gene *fushi tarazu (ftz)*, clusters of silver grains are first detectable over the nuclei during blastoderm formation.

At first, the H42 gene becomes active in all nuclei, as the nuclear labelling indicates, but subsequently its transcripts accumulate in a blastodermal belt. Within the belt the labelling is cytoplasmic as well as nuclear. The appearance of this belt could be due to other genes repressing the release of H42 transcripts into the cytoplasm in certain regions and stages. Alternatively, the transcript may be degraded immediately outside the nuclei but in the belt region the rate of transcription may exceed that of degradation. The number of silver grains is much higher over the cells of the belt (cytoplasm and nuclei) than over those in the other regions of the embryo (nuclei only), and this would support the second assumption.

Since our probe contains the M repeat 3' to the homeobox (Fig. 3), we cannot rule out some non-specific labelling due to cross-hybridization with other M-repeat sequences. In this case, however, we would expect a more or less irregular pattern or higher background. The regularity and intensity of the signals over the nuclei as well as the distinctness and high intensity of the cytoplasmic labelling in a region and developmental stage comparable with the Dfd gene's transcription in the *Drosophila* embryo argues against this possibility.

Acknowledgements. We thank P. LeMotte for critical reading of the manuscript. These investigations were supported by the Deutsche Forschungsgemeinschaft (F1 151/1-1 and 1-2; Wa 556/1-1 and 1-2)

References

- Boncinelli E, Simone A, LaVolpe A, Faiella A, Fidauza V, Acampora D, Scotto L (1985) Human cDNA clones containing homeobox sequences. Cold Spring Harbor Symp Quant Biol 50:301-306
- Campos-Ortega J, Hartenstein V (1985) Embryonic development of *Drosophila melanogaster*. Springer Verlag, Berlin
- Carrasco AE, McGinnis W, Gehring WJ, DeRobertis EM (1984) Cloning of a X. laevis gene expressed during early embryogenesis coding for a peptide region homologous to Drosophila homeotic genes. Cell 37:409–414
- Chadwick R, McGinnis W (1987) Temporal and spatial distribution of transcripts from the *Deformed* gene of *Drosophila*. EMBO J 6:779-789
- Colberg-Poly AM, Voss SD, Chowdhury K, Gruss P (1985) Structural analysis of murine genes containing homeobox sequences and their expression in embryonal carcinoma cells. Nature 314:713-718
- Desplan C, Theis J, O'Farrell PH (1985) The Drosophila developmental gene, engrailed, encodes a sequence-specific DNA binding activity. Nature 318:630–635
- Du Praw EJ (1967) The honeybee embryo. In: Wilt FH, Wessells NK (eds) Methods in Developmental Biology. Th Y Crowell Co, New York, pp 183–217

- Falzon M, Sanderson N, Chung SY (1987) Cloning and expression of rat homeobox-containing sequences. Gene 54:23–32
- Fleig R, Sander K (1985) Blastoderm development in honey bee embryogenesis as seen in the scanning electron microscope. Int J Invert Reprod Dev 8:279-286
- Fleig R, Sander K (1986) The embryogenesis of the honeybee Apis mellifera L. (Hymenoptera: Apidae): a SEM study. Int J Insect Morphol Embryol 15:449–462
- Frischauf AM, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying polylinker sequences. J Mol Biol 170:827–842
- Gehring WJ (1987) Homeotic genes, the homeobox, and the spatial organization of the embryo. Harvey Lect 81:153–172
- Hafen E, Levine M, Garber RL, Gehring WJ (1983) An improved in situ hybridization method for the detection of cellular RNAs in *Drosophila* tissue sections and its application for localizing transcripts of the homeotic *Antennapedia* gene complex. EMBO J 2:617-623
- Hafen E, Kuroiwa A, Gehring WJ (1984) Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Dro-sophila* embryonic development. Cell 37:833–841
- Harding K, Wedeen C, McGinnis W, Levine M (1985) Spatially regulated expression of homeotic genes in *Drosophila*. Science 229:1236–1242
- Joyner AL, Martin GR (1987) En-1 and En-2, two mouse genes with sequence homology to the *Drosophila engrailed* gene: expression during embryogenesis. Genes Dev 1:29–38
- Krause G (1939) Die Eitypen der Insekten. Biol Zbl 59:495-536
- Laughon A, Scott MP (1984) Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature 310:25–31
- Levine M, Rubin GN, Tjian R (1984) Human DNA sequences homologous to a protein-coding region conserved between homeotic genes of *Drosophila*. Cell 38:667–678
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, New York
- Martinez-Arias A, Ingham PW, Scott MP, Akam ME (1987) The spatial and temporal development of *Dfd* and *Scr* transcripts throughout development of *Drosophila*. Dev 100:673–683
- McGinnis W, Levine M, Hafen E, Kuroiwa A, Gehring WJ (1984a) A conserved DNA sequence in homeotic genes of the *Drosophila Antennapedia* and *bithorax* complexes. Nature 308:428–433
- McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring WJ (1984b) A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. Cell 37:403–408
- McGinnis W, Hart CP, Gehring WJ, Ruddle FH (1984c) Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of *Drosophila*. Cell 38:675-680
- Messing J, Vieira J (1982) A new pair of M13 vectors for selecting

either DNA strand of double-digest restriction fragments. Gene 19:269–276

- Morgan TH, Bridges CB, Sturtevant AH (1925) The genetics of Drosophila. Martinus Nijhoff, Gravenhage
- Müller MM, Carrasco AE, DeRobertis EM (1984) A homeo-boxcontaining gene expressed during oogenesis in *Xenopus*. Cell 39:157–162
- Nelson JA (1915) The embryology of the honey bee. Princeton University Press, Princeton New York
- Nüsslein-Volhard C, Wieschaus E, Jürgens G (1982) Segmentierung bei Drosophila – eine genetische Analyse. Verh Dtsch Zool Ges 1982:91–104
- Regulski M, McGinnis N, Chadwick R, McGinnis W (1987) Developmental and molecular analysis of *Deformed*; a homeotic gene controlling *Drosophila* head development. EMBO J 6:767–777
- Sander K (1976) Specification of the basic body pattern in insect embryogenesis. Adv Insect Physiol 12:125–238
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463– 5467
- Schnetter M (1935) Morphologische Untersuchungen über das Differenzierungszentrum in der Embryonalentwicklung der Honigbiene. Z Morphol Ökol Tiere 29:114–195
- Scott MP, Weiner AJ (1984) Structural relationship among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. Proc Natl Acad Sci USA 81:4115–4119
- Shepherd JCW, McGinnis W, Carrasco AE, DeRobertis EM, Gehring WJ (1984) Fly and frog homeo domains show homologies with yeast mating type regulatory proteins. Nature 310:70– 71
- Turner FR, Mahowald AP (1976) Scanning electron microscopy of *Drosophila melanogaster*. I. The structure of the egg envelopes and the formation of the cellular blastoderm. Dev Biol 50:95–108
- Walldorf U, Richter S, Ryseck RP, Steller H, Edström JE, Bautz EKF, Hovemann B (1984) Cloning of heat shock locus 93D from *Drosophila melanogaster*. EMBO J 3:2499–2504
- Weir MP, Kornberg T (1985) Patterns of *engrailed* and *fushi tarazu* transcripts reveal novel intermediate stages in *Drosophila* segmentation. Nature 318:433–439
- Wharton KA, Yedvobnick B, Finnerty VG, Artavanis-Tsakonas S (1985) A novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in Drosophila melanogaster. Cell 40:55–62
- Zalokar M, Erk J (1977) Phase-partition fixation and staining of Drosophila eggs. Stain Technol 52:89–95

Received March 18, 1988 Accepted in revised form May 30, 1988